

С.Г. СЕЛЕВИЧ, канд. техн. наук, доц. НТУ «ХПИ»;
С.А. ПИСМЕНИЦКИЙ, студент НТУ «ХПИ»;
В.А. ЧАЛЫЙ, студент НТУ «ХПИ»

ОЦЕНКА ДОСТОВЕРНОСТИ ИНФОРМАЦИИ О СКОРОСТЯХ ДВИЖЕНИЯ ТРАНСПОРТНЫХ ПОТОКОВ СИСТЕМОЙ АГРЕГАЦИИ ДАННЫХ АВТОМОБИЛЬНЫХ GPS НАВИГАТОРОВ

В статье выполнен статистический анализ ошибки систем агрегации данных GPS навигаторов.

У статті виконаний статистичний аналіз помилки систем агрегації даних GPS навігаторів.

The statistical analysis of GPS-based systems accuracy was made.

Введение

Результаты измерения скоростей транспортного потока и оценки времени проезда находят все большее применение на практике. Так измерения в режиме реального времени являются основой для современных информационных систем поддержки водителя (ATIS), которые, в свою очередь, позволяют осуществлять осознанный выбор маршрута движения, времени выезда и др.

Наиболее распространенным подходом сбора исходных данных для ATIS является установка локальных детекторов и видеокамер, выполняющих измерения на конкретном участке дорожной сети [1,2]. В силу инсталляционных и стоимостных ограничений локальные детекторы устанавливаются дискретно на небольших участках сети и, как следствие, не способны предоставить большое покрытие. Альтернативным способом сбора данных о транспортных потоках является использование GPS навигаторов, специально установленных на коммерческом автотранспорте [3] или находящихся в составе мобильных телефонов [4]. Это позволяет, задействовав существующие сотовые сети для передачи данных, существенно снизить инсталляционные затраты, обеспечить высокую степень покрытия и исключить ошибки в измерении скорости и местонахождения. При этом достоверность таких данных зависит от целого ряда параметров, наиболее значимыми из которых являются: механизм агрегации и выдачи данных, количество активных GPS навигаторов на участке дорожной сети, погрешность определения координат GPS навигатором.

Цель работы и постановка задачи

Целью данной работы является получение статистических оценок, их доверительных интервалов, а также выбор функции распределения для

модуля ошибки скорости движения транспортного потока, предоставляемой общедоступной системой [4].

Порядок измерений и первичная обработка

В ходе выполнения эксперимента были получены скорости движения транспортного потока на 27 участках улично-дорожной сети одновременно двумя способами: путем непосредственной регистрации данных предоставляемых системой [4] $x_{grandex}$ и посредством косвенного измерения скорости движения транспортного потока $x_{изм}$ по методике [5]. Тогда величина рассогласования:

$$x_i = |x_{i,изм} - x_{i,grandex}| \quad (1)$$

представляет собой модуль абсолютной ошибки измерения скорости движения транспортного потока на i-ом участке улично-дорожной сети.

В процессе первичной обработки были исключены результаты 13-го опыта, поскольку они соответствуют режиму движения «свободный поток», в то время как, методика вычисления $x_{изм}$ рассчитана на режим «синхронизированный поток», когда водители теряют возможность маневрировать и вынуждены согласовывать свою скорость со скоростью потока [6].

Определение выборочных оценок и доверительных интервалов

Статистическая обработка полученных данных была выполнена по методике, представленной в [7]. Так выборочные математическое ожидание \tilde{m}_x , дисперсия \tilde{D}_x , среднеквадратичное отклонение $\tilde{\sigma}_x$, асимметрия \tilde{a}_x и эксцесс \tilde{g}_x были вычислены по формулам:

$$\tilde{m}_x = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

$$\tilde{D}_x = \frac{1}{n-1} \sum_{i=1}^n (x_i - \tilde{m}_x)^2 \quad (3)$$

$$\tilde{\sigma}_x = \sqrt{\tilde{D}_x} \quad (4)$$

$$\tilde{a}_x = \frac{1}{(n-1)\tilde{\sigma}_x^3} \sum_{i=1}^n (x_i - \tilde{m}_x)^3 \quad (5)$$

$$\tilde{g}_x = \frac{1}{(n-1)\tilde{\sigma}_x^4} \sum_{i=1}^n (x_i - \tilde{m}_x)^4 - 3 \quad (6)$$

Результаты вычисления выборочных параметров распределения приведены в таблице 1.

Таблица 1 – Выборочные параметры распределения

Параметр	\tilde{m}_x	\tilde{D}_x	$\tilde{\sigma}_x$	\tilde{a}_x	\tilde{e}_x
Величина	9,31	59,26	7,70	0,76	-0,43

По выборочным параметрам были определены доверительные интервалы для генеральных параметров. Доверительный интервал для генерального математического ожидания m_x :

$$\tilde{m}_x - \frac{\tilde{\sigma}_x t_{1-\frac{\alpha}{2}}(f)}{\sqrt{n}} \leq m_x \leq \tilde{m}_x + \frac{\tilde{\sigma}_x t_{1-\frac{\alpha}{2}}(f)}{\sqrt{n}} \quad (7)$$

где $t_p(f)$ – квантиль t -распределения Стьюдента, соответствующий вероятности p , n – объём выборки, $f = n-1$ – число степеней свободы выборки.

Доверительный интервал для генеральной дисперсии D_x :

$$\frac{f \tilde{D}_x}{\chi^2_{1+\frac{\alpha}{2}}(f)} \leq D_x \leq \frac{f \tilde{D}_x}{\chi^2_{1-\frac{\alpha}{2}}(f)} \quad (8)$$

где $\chi_p^2(f)$ – квантиль χ^2 -распределения Пирсона.

Формулы (7-8) имеют место, только если исходная величина X распределена по нормальному закону. Поэтому результаты, полученные по этим формулам, являются приближёнными.

Доверительный интервал для генеральных асимметрии a_x и эксцесса e_x , с использованием неравенства Чебышёва, вычислим по формулам:

$$\tilde{a}_x - \sqrt{\frac{D_a}{q}} \leq a_x \leq \tilde{a}_x + \sqrt{\frac{D_a}{q}}, \quad (9)$$

$$\tilde{e}_x - \sqrt{\frac{n_e}{q}} \leq e_x \leq \tilde{e}_x + \sqrt{\frac{n_e}{q}}, \quad (10)$$

где $q=1-p$ – уровень значимости, D_a , D_e – дисперсии выборочных асимметрии и эксцесса, которые определяются по формулам:

$$D_a = \frac{6(n-1)}{(n+1)(n+3)}, \quad (11)$$

$$D_e = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}. \quad (12)$$

Результаты вычисления доверительных интервалов для доверительной вероятности $p = 0.9$, числа степеней свободы выборки $f = 25$, дисперсии асимметрии $D_a = 0.9$ и дисперсии эксцесса $D_e = 0.53$ представлены в таблице 2.

Таблица 2 – Доверительные интервалы генеральных параметров

Параметр	m_x	D_x	a_x	e_x
Доверительный интервал	[6,72 .. 11,89]	[39,35 .. 101,40]	[-0,62 .. 2,15]	[-2,72 .. 1,86]

Выбор функции распределения

Для выдвижения гипотезы о виде теоретического распределения построим гистограмму. Выбор числа интервалов k для построения гистограммы осуществим по формулы Стёрджесса:

$$k = [1 + 3.322 \ln(n)], \quad (1)$$

3)

На рис. 1 изображена гистограмма для $k=6$ и кривые некоторых двух- и однопараметрических законов распределения, полученные с использованием принципа максимального правдоподобия.

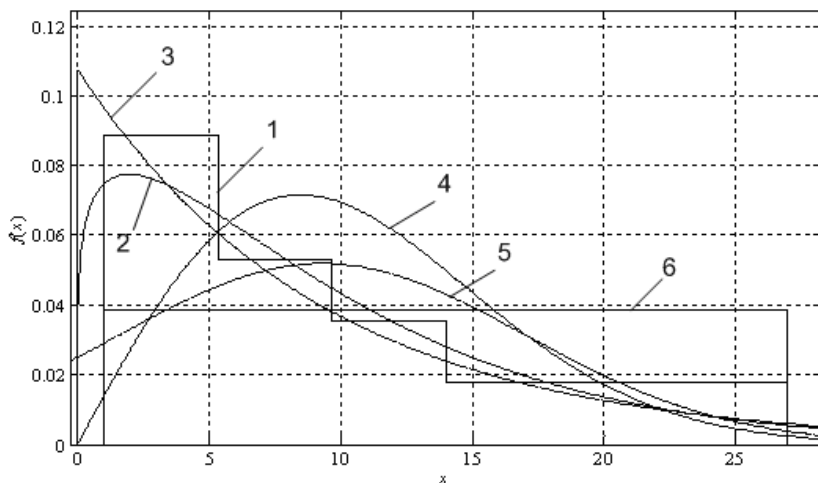


Рисунок 1 – Гистограмма и кривые теоретического распределения случайной величины X : 1 – гистограмма распределения случайной величины X ; 2 – распределение Вейбулла с параметрами $\lambda = 9.31$ и $k = 7.70$; 3 – показательное распределение с параметром $\alpha = 0.11$; 4 – нормальное распределение с параметрами

$m = 9.31$ и $\sigma = 7.70$; 5 – Рэлеевское распределение с параметром $\sigma = 8.47$; 6 – равномерное распределение с параметрами $a = 1.00$ и $b = 27.00$

Для обоснования выбора теоретического распределения использован критерий согласия Колмогорова. Так максимальное значение критического уровня значимости соответствует распределению Вейбулла и составляет $q = 0.94$.

Вывод

Полученные в статье точечные оценки модуля ошибки и их доверительные интервалы позволяют говорить о достаточной для дальнейших исследований в составе макромоделей степени точности. Обоснование выбора функции распределения в виде распределения Вейбулла позволяет учесть ошибку при построении макромоделей.

Список литературы: 1. S. Lämmer and D. Helbing (2010) Self-Stabilizing Decentralized Signal Control of Realistic, Saturated Network Traffic . Santa Fe Working Paper Nr. 10-09-019. 2. «Видеопробки» [Электронный ресурс] // <http://videoprobki.ua/>. 3. Reinhart, K., Schafer, R.P., 2006. Taxi FCD – jam detection by means of floating car data (FCD), <http://www.dlr.de/vf/forschung/projekte/taxi_fcd> (accessed 12.11.06) 4. «Yandex.Пробки» [Электронный ресурс] // «Яндекс». http://company.yandex.ru/researches/reports/ya_bottlenecks_kiev_11.xml/. 5. Edie L., 1965. Discussion on traffic stream measurements and definitions. In: Proceedings of the Second International Symposium on the Theory of Traffic Flow. Paris, France, pp. 139–154. 6. Kerner B.S. Three-Phase Traffic Theory and Highway Capacity, Networks and Spatial Economics, 2000, v.2, p.46-56. 7. Иглин С.И. Математические расчеты на базе MATLAB. – СПб.: БХВ-Петербург, 2005. – 640с.: ил.

Поступила в редакцию 11.04.2012